SEQUENCE LISTING

<110> Astex Technology Limited
Cosme, Jose
Ward, Alison
Vuillard, Laurent
Williams, Pamela
Hamilton, Bruce

<120> Methods of Purification of Cytochrome P450 Proteins

<130> AHBCP6047252

<140> PCT/GB02/02668
<141> 2002-05-30

<160> 84

<170> PatentIn Ver. 2.1

<210> 1
<211> 1428
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C19 (internal deletion, and His tagged) coding sequence.

<400> 1
atggctaaga aaacgagctc taaagggcgg ccgcctggcc ccactcctct cccagtgatt 60
ggaaatatcc tacagataga tattaaggat gtcagcaaat ccttaaccaa tctctcaaaa 120
atctatggcc ctgtgttcac tctgtatttt ggcctggaac gcatggtggt gctgcatgga 180
tatgaagtgg tgaaggaagc cctgattgat cttggagagg agttttctgg aagaggccat 240
ttcccactgg ctgaaagagc taacagagga tttggaatcg ttttcagcaa tggaaagaga 300
tggaaggaga tccggcgttt ctccctcatg acgctgcgga attttgggat ggggaagagg 360
agcattgagg accgtgttca agaggaagcc cactgccttg tggaggagtt gagaaaaacc 420
aaggcttcac cctgtgatcc cactttcatc ctgggctgtg ctccctgcaa tgtgatctgc 480
tccattattt tccagaaacg tttcgattat aaagatcagc aatttcttaa cttgatggaa 540
aaattgaatg aaaacatcag gattgtaagc acccccctgga tccagatatg caataatttt 600
cccactatca ttgattattt cccgggaacc cataacaaat tacttaaaaa ccttgctttt 660
atggaaagtg atattttgga gaaagtaaaa gaacaccaag aatcgatgga catcaacaac 720
cctcgggact ttattgattg cttcctgatc aaaatggaga aggaaaagca aaaccaacag 780
tctgaattca ctattgaaaa cttggtaatc actgcagctg acttacttgg agctgggaca 840
gagacaacaa gcacaaccct gagatatgct ctccttctcc tgctgaagca cccagaggtc 900
acagctaaag tccaggaaga gattgaacgt gtcgttggca gaaaccggag cccctgcatg 960
caggacaggg gccacatgcc ctacacagat gctgtggtgc acgaggtcca gagatacatc 1020
gacctcatcc ccaccagcct gccccatgca gtgacctgtg acgttaaatt cagaaactac 1080
ctcattccca agggcacaac catattaact tccctcactt ctgtgctaca tgacaacaaa 1140
gaatttccca acccagagat gtttgaccct cgtcactttc tgcatgaagg tggaaatttt 1200
aagaaaagta actacttcat gcctttctca gcaggaaaac ggatttgtgt gggagagggc 1260
ctggcccgca tggagctgtt tttattcctg accttcattt tacagaactt taacctgaaa 1320
tctctgattg acccaaagga ccttgacaca actcctgttg tcaatggatt tgcttctgtc 1380
ccgcccttct accagctctg cttcattcct gtccaccacc accactga 1428

<210> 2

<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Protein
sequence of 2C19 coded by SEQ ID NO: 1

<400> 2

```
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
  1               5                  10                  15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Asp Ile Lys Asp Val Ser
             20                  25                  30

Lys Ser Leu Thr Asn Leu Ser Lys Ile Tyr Gly Pro Val Phe Thr Leu
         35                  40                  45

Tyr Phe Gly Leu Glu Arg Met Val Val Leu His Gly Tyr Glu Val Val
     50                  55                  60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly His
 65                  70                  75                  80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser
                 85                  90                  95

Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
            100                 105                 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
        115                 120                 125

Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
    130                 135                 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145                 150                 155                 160

Ser Ile Ile Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
                165                 170                 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Arg Ile Val Ser Thr Pro
            180                 185                 190

Trp Ile Gln Ile Cys Asn Asn Phe Pro Thr Ile Ile Asp Tyr Phe Pro
        195                 200                 205

Gly Thr His Asn Lys Leu Leu Lys Asn Leu Ala Phe Met Glu Ser Asp
    210                 215                 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Ile Asn Asn
225                 230                 235                 240

Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu Lys Glu Lys
                245                 250                 255
```

Gln Asn Gln Gln Ser Glu Phe Thr Ile Glu Asn Leu Val Ile Thr Ala
260 265 270

Ala Asp Leu Leu Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
275 280 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
290 295 300

Gln Glu Glu Ile Glu Arg Val Val Gly Arg Asn Arg Ser Pro Cys Met
305 310 315 320

Gln Asp Arg Gly His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
325 330 335

Gln Arg Tyr Ile Asp Leu Ile Pro Thr Ser Leu Pro His Ala Val Thr
340 345 350

Cys Asp Val Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
355 360 365

Leu Thr Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
370 375 380

Pro Glu Met Phe Asp Pro Arg His Phe Leu His Glu Gly Gly Asn Phe
385 390 395 400

Lys Lys Ser Asn Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe Leu Thr Phe
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Ile Asp Pro Lys Asp Leu
435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

Gln Leu Cys Phe Ile Pro Val His His His His
465 470 475

<210> 3
<211> 1428
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C19 wild type
1B

<400> 3
atggctaaga aaacgagctc taaagggcgg ccgcctggcc ctactcctct cccagtgatt 60
ggaaatatcc tacagataga tattaaggat gtcagcaaat ccttaaccaa tctctcaaaa 120
atctatggcc ctgtgttcac tctgtatttt ggcctggaac gcatggtggt gctgcatgga 180
tatgaagtgg tgaaggaagc cctgattgat cttggagagg agttttctgg aagaggccat 240

```
ttcccactgg ctgaaagagc taacagagga tttggaatcg ttttcagcaa tggaaagaga 300
tggaaggaga tccggcgttt ctccctcatg acgctgcgga attttgggat ggggaagagg 360
agcattgagg accgtgttca agaggaagcc cgctgccttg tggaggagtt gagaaaaacc 420
aaagcttcac cctgtgatcc cactttcatc ctgggctgtg ctccctgcaa tgtgatctgc 480
tccattattt tccagaaacg tttcgattat aaagatcagc aatttcttaa cttgatggaa 540
aaattgaatg aaaacatcag gattgtaagc acccccctgga tccagatatg caataatttt 600
cccactatca ttgattattt cccgggaacc cataacaaat tacttaaaaa ccttgctttt 660
atggaaagtg atattttgga gaaagtaaaa gaacaccaag aatcgatgga catcaacaac 720
cctcgggact ttattgattg cttcctgatc aaaatggaga aggaaaagca aaaccaacag 780
tctgaattca ctattgaaaa cttggtaatc actgcagctg acttacttgg agctgggaca 840
gagacaacaa gcacaaccct gagatatgct ctccttctcc tgctgaagca cccagaggtc 900
acagctaaag tccaggaaga gattgaacgt gtcgttggca gaaaccggag cccctgcatg 960
caggacaggg gccacatgcc ctacacagat gctgtggtgc acgaggtcca gagatacatc 1020
gacctcatcc ccaccagcct gccccatgca gtgacctgtg acgttaaatt cagaaactac 1080
ctcattccca agggcacaac catattaact tccctcactt ctgtgctaca tgacaacaaa 1140
gaatttccca acccagagat gtttgaccct cgtcactttc tggatgaagg tggaaatttt 1200
aagaaaagta actacttcat gcctttctca gcaggaaaac ggatttgtgt gggagagggc 1260
ctggcccgca tggagctgtt tttattcctg accttcattt tacagaactt taacctgaaa 1320
tctctgattg acccaaagga ccttgacaca actcctgttg tcaatggatt tgcttctgtc 1380
ccgcccttct accagctctg cttcattcct gtccaccacc accactga              1428
```

<210> 4
<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Translation of
SEQ ID NO:3

<400> 4

```
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
  1               5                  10                  15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Asp Ile Lys Asp Val Ser
             20                  25                  30

Lys Ser Leu Thr Asn Leu Ser Lys Ile Tyr Gly Pro Val Phe Thr Leu
         35                  40                  45

Tyr Phe Gly Leu Glu Arg Met Val Val Leu His Gly Tyr Glu Val Val
     50                  55                  60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly His
 65                  70                  75                  80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser
                 85                  90                  95

Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
            100                 105                 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
        115                 120                 125

Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
```

130 135 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145 150 155 160

Ser Ile Ile Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
165 170 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Arg Ile Val Ser Thr Pro
180 185 190

Trp Ile Gln Ile Cys Asn Asn Phe Pro Thr Ile Ile Asp Tyr Phe Pro
195 200 205

Gly Thr His Asn Lys Leu Leu Lys Asn Leu Ala Phe Met Glu Ser Asp
210 215 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Ile Asn Asn
225 230 235 240

Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu Lys Glu Lys
245 250 255

Gln Asn Gln Gln Ser Glu Phe Thr Ile Glu Asn Leu Val Ile Thr Ala
260 265 270

Ala Asp Leu Leu Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
275 280 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
290 295 300

Gln Glu Glu Ile Glu Arg Val Val Gly Arg Asn Arg Ser Pro Cys Met
305 310 315 320

Gln Asp Arg Gly His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
325 330 335

Gln Arg Tyr Ile Asp Leu Ile Pro Thr Ser Leu Pro His Ala Val Thr
340 345 350

Cys Asp Val Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
355 360 365

Leu Thr Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
370 375 380

Pro Glu Met Phe Asp Pro Arg His Phe Leu Asp Glu Gly Gly Asn Phe
385 390 395 400

Lys Lys Ser Asn Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe Leu Thr Phe
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Ile Asp Pro Lys Asp Leu

435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

Gln Leu Cys Phe Ile Pro Val His His His His
465 470 475

<210> 5
<211> 1443
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2D6 encoding nucleic acid

<400> 5
atggctaaaa aaacctcttc taaaggccga ccgccgggtc cgctgccgct gccaggcctg 60
ggtaacctgc tgcatgtgga cttccagaac accccgtact gcttcgacca gctgcgtcgt 120
cgtttcggtg acgtgttctc tctgcagctg gcttggaccc cggttgttgt tctgaacggt 180
ctggctgctg ttcgcgaagc tctggttacc cacggtgaag acaccgctga ccgtccgccg 240
gtcccgatca cccagatcct gggttttggt ccgcgttccc aaggtgtttt cctggctcgt 300
tacggaccgg cttggcgtga acagcgtcgt ttctctgttt ctaccctgcg taacctgggt 360
ctgggtaaaa aatctctgga acagtgggtt accgaagaag ctgcatgcct gtgcgctgct 420
ttcgctaacc actctggtcg tccgttccgt ccgaacggtc tgctggacaa agctgtttct 480
aacgttatcg cttctctgac ctgcggccgc cgtttcgaat acgacgaccc gcgtttcctg 540
cgtctgctgg acctggctca ggaaggtctg aaagaggagt ctggtttcct gcgtgaagtt 600
ctgaacgctg ttccggttct gctgcacatc ccagctctgg ctggtaaagt tctgcgtttc 660
cagaaagcat tcctgaccca gctggacgaa ctgctgaccg aacaccgtat gacctgggac 720
ccggctcagc cgccacgtga cctgaccgaa gctttcctgg ctgaaatgga aaaagctaaa 780
ggtaacccgg aatcttcttt caacgatgaa aatctgcgta tcgttgttgc tgacctgttc 840
tccgcgggta tggttaccac ctctaccacc ctggcttggg gtctgctgct gatgatcctg 900
cacccggatg tacagcgtcg tgttcagcag gaaatcgacg acgttattgg ccaggttcgt 960
cggccggaaa tgggtgacca ggctcacatg ccgtacacca ccgctgttat ccacgaagtt 1020
cagcgcttcg gtgacatcgt tccgctgggt atgacccaca tgacctctcg tgacatcgaa 1080
gttcagggtt tccgtatccc gaaaggtacc accctgatca ccaacctgtc ttctgttctg 1140
aaagacgaag ctgtttggga aaaaccgttc cgtttccatc cggaacactt cctggacgct 1200
cagggtcact tcgttaaacc ggaagctttc ctgccgttct ctgctggtcg tcgtgcttgc 1260
ctgggtgaac cgctggctcg tatggaactg ttcctgttct tcacctctct gctgcagcac 1320
ttctctttct ctgttccgac cggtcagccg cgtccgtctc accacggtgt tttcgctttc 1380
ctggtttctc cgtctccgta cgaactgtgc gctgttccgc gtggagctca ccaccaccac 1440
tga 1443

<210> 6
<211> 480
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Translation of SEQ ID NO: 5

<400> 6
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Leu Pro

```
1               5                  10                  15

Leu Pro Gly Leu Gly Asn Leu Leu His Val Asp Phe Gln Asn Thr Pro
            20                  25                  30

Tyr Cys Phe Asp Gln Leu Arg Arg Arg Phe Gly Asp Val Phe Ser Leu
        35                  40                  45

Gln Leu Ala Trp Thr Pro Val Val Val Leu Asn Gly Leu Ala Ala Val
    50                  55                  60

Arg Glu Ala Leu Val Thr His Gly Glu Asp Thr Ala Asp Arg Pro Pro
65                  70                  75                  80

Val Pro Ile Thr Gln Ile Leu Gly Phe Gly Pro Arg Ser Gln Gly Val
                85                  90                  95

Phe Leu Ala Arg Tyr Gly Pro Ala Trp Arg Glu Gln Arg Arg Phe Ser
            100                 105                 110

Val Ser Thr Leu Arg Asn Leu Gly Leu Gly Lys Lys Ser Leu Glu Gln
        115                 120                 125

Trp Val Thr Glu Glu Ala Ala Cys Leu Cys Ala Ala Phe Ala Asn His
    130                 135                 140

Ser Gly Arg Pro Phe Arg Pro Asn Gly Leu Leu Asp Lys Ala Val Ser
145                 150                 155                 160

Asn Val Ile Ala Ser Leu Thr Cys Gly Arg Arg Phe Glu Tyr Asp Asp
                165                 170                 175

Pro Arg Phe Leu Arg Leu Leu Asp Leu Ala Gln Glu Gly Leu Lys Glu
            180                 185                 190

Glu Ser Gly Phe Leu Arg Glu Val Leu Asn Ala Val Pro Val Leu Leu
        195                 200                 205

His Ile Pro Ala Leu Ala Gly Lys Val Leu Arg Phe Gln Lys Ala Phe
    210                 215                 220

Leu Thr Gln Leu Asp Glu Leu Leu Thr Glu His Arg Met Thr Trp Asp
225                 230                 235                 240

Pro Ala Gln Pro Pro Arg Asp Leu Thr Glu Ala Phe Leu Ala Glu Met
                245                 250                 255

Glu Lys Ala Lys Gly Asn Pro Glu Ser Ser Phe Asn Asp Glu Asn Leu
            260                 265                 270

Arg Ile Val Val Ala Asp Leu Phe Ser Ala Gly Met Val Thr Thr Ser
        275                 280                 285

Thr Thr Leu Ala Trp Gly Leu Leu Leu Met Ile Leu His Pro Asp Val
    290                 295                 300

Gln Arg Arg Val Gln Gln Glu Ile Asp Asp Val Ile Gly Gln Val Arg
```

```
305                 310                 315                 320

Arg Pro Glu Met Gly Asp Gln Ala His Met Pro Tyr Thr Thr Ala Val
                325                 330                 335

Ile His Glu Val Gln Arg Phe Gly Asp Ile Val Pro Leu Gly Met Thr
            340                 345                 350

His Met Thr Ser Arg Asp Ile Glu Val Gln Gly Phe Arg Ile Pro Lys
        355                 360                 365

Gly Thr Thr Leu Ile Thr Asn Leu Ser Ser Val Leu Lys Asp Glu Ala
    370                 375                 380

Val Trp Glu Lys Pro Phe Arg Phe His Pro Glu His Phe Leu Asp Ala
385                 390                 395                 400

Gln Gly His Phe Val Lys Pro Glu Ala Phe Leu Pro Phe Ser Ala Gly
                405                 410                 415

Arg Arg Ala Cys Leu Gly Glu Pro Leu Ala Arg Met Glu Leu Phe Leu
            420                 425                 430

Phe Phe Thr Ser Leu Leu Gln His Phe Ser Phe Ser Val Pro Thr Gly
        435                 440                 445

Gln Pro Arg Pro Ser His His Gly Val Phe Ala Phe Leu Val Ser Pro
    450                 455                 460

Ser Pro Tyr Glu Leu Cys Ala Val Pro Arg Gly Ala His His His His
465                 470                 475                 480
```

<210> 7
<211> 1458
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 3A4

<400> 7
atggcatacg gtactcattc acatggtctg tttaaaaaac tgggaattcc agggcccaca 60
cctctgcctt ttttgggaaa tattttgtcc taccataagg gcttttgtat gtttgacatg 120
gaatgtcata aaaagtatgg aaaagtgtgg ggcttttatg atggtcaaca gcctgtgctg 180
gctatcacag atcctgacat gatcaaaaca gtgctagtga aagaatgtta ttctgtcttc 240
acaaaccgga ggcctttggg tccagtggga tttatgaaaa gtgccatctc tatagctgag 300
gatgaagaat ggaagagatt acgatcattg ctgtctccaa ccttcaccag tggaaaactc 360
aaggagatgg tccctatcat tgcccagtat ggagatgtgt tggtgagaaa tctgaggcgg 420
gaagcagaga caggcaagcc tgtcaccttg aaagacgtct ttggggccta cagcatggat 480
gtgatcacta gcacatcatt tggagtgaac atcgactctc tcaacaatcc acaagacccc 540
tttgtggaaa acaccaagaa gcttttaaga tttgattttt tggatccatt ctttctctca 600
ataacagtct ttccattcct catcccaatt cttgaagtat taaatatctg tgtgtttcca 660
agagaagtta caaatttttt aagaaaatct gtaaaaagga tgaaagaaag tcgcctcgaa 720

```
gatacacaaa agcaccgagt ggatttcctt cagctgatga ttgactctca gaattcaaaa 780
gaaactgagt cccacaaagc tctgtccgat ctggagctcg tggcccaatc aattatcttt 840
atttttgctg gctatgaaac cacgagcagt gttctctcct tcattatgta tgaactggcc 900
actcaccctg atgtccagca gaaactgcag gaggaaattg atgcagtttt acccaataag 960
gcaccaccca cctatgatac tgtgctacag atggagtatc ttgacatggt ggtgaatgaa 1020
acgctcagat tattcccaat tgctatgaga cttgagaggg tctgcaaaaa agatgttgag 1080
atcaatggga tgttcattcc caaaggggtg gtggtgatga ttccaagcta tgctcttcac 1140
cgtgacccaa agtactggac agagcctgag aagttcctcc ctgaaagatt cagcaagaag 1200
aacaaggaca acatagatcc ttacatatac acacccttg gaagtggacc cagaaactgc 1260
attggcatga ggtttgctct catgaacatg aaacttgctc taatcagagt ccttcagaac 1320
ttctccttca aaccttgtaa agaaacacag atccccctga aattaagctt aggaggactt 1380
cttcaaccag aaaaacccgt tgttctaaag gttgagtcaa gggatggcac cgtaagtgga 1440
gcccaccatc accattga                                              1458
```

<210> 8
<211> 485
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 3A4

<400> 8

```
Met Ala Tyr Gly Thr His Ser His Gly Leu Phe Lys Lys Leu Gly Ile
  1               5                  10                  15

Pro Gly Pro Thr Pro Leu Pro Phe Leu Gly Asn Ile Leu Ser Tyr His
             20                  25                  30

Lys Gly Phe Cys Met Phe Asp Met Glu Cys His Lys Lys Tyr Gly Lys
         35                  40                  45

Val Trp Gly Phe Tyr Asp Gly Gln Gln Pro Val Leu Ala Ile Thr Asp
     50                  55                  60

Pro Asp Met Ile Lys Thr Val Leu Val Lys Glu Cys Tyr Ser Val Phe
 65                  70                  75                  80

Thr Asn Arg Arg Pro Phe Gly Pro Val Gly Phe Met Lys Ser Ala Ile
                 85                  90                  95

Ser Ile Ala Glu Asp Glu Glu Trp Lys Arg Leu Arg Ser Leu Leu Ser
            100                 105                 110

Pro Thr Phe Thr Ser Gly Lys Leu Lys Glu Met Val Pro Ile Ile Ala
        115                 120                 125

Gln Tyr Gly Asp Val Leu Val Arg Asn Leu Arg Arg Glu Ala Glu Thr
    130                 135                 140

Gly Lys Pro Val Thr Leu Lys Asp Val Phe Gly Ala Tyr Ser Met Asp
145                 150                 155                 160

Val Ile Thr Ser Thr Ser Phe Gly Val Asn Ile Asp Ser Leu Asn Asn
                165                 170                 175
```

```
Pro Gln Asp Pro Phe Val Glu Asn Thr Lys Lys Leu Leu Arg Phe Asp
            180                 185                 190

Phe Leu Asp Pro Phe Phe Leu Ser Ile Thr Val Phe Pro Phe Leu Ile
        195                 200                 205

Pro Ile Leu Glu Val Leu Asn Ile Cys Val Phe Pro Arg Glu Val Thr
    210                 215                 220

Asn Phe Leu Arg Lys Ser Val Lys Arg Met Lys Glu Ser Arg Leu Glu
225                 230                 235                 240

Asp Thr Gln Lys His Arg Val Asp Phe Leu Gln Leu Met Ile Asp Ser
                245                 250                 255

Gln Asn Ser Lys Glu Thr Glu Ser His Lys Ala Leu Ser Asp Leu Glu
            260                 265                 270

Leu Val Ala Gln Ser Ile Ile Phe Ile Phe Ala Gly Tyr Glu Thr Thr
        275                 280                 285

Ser Ser Val Leu Ser Phe Ile Met Tyr Glu Leu Ala Thr His Pro Asp
    290                 295                 300

Val Gln Gln Lys Leu Gln Glu Glu Ile Asp Ala Val Leu Pro Asn Lys
305                 310                 315                 320

Ala Pro Pro Thr Tyr Asp Thr Val Leu Gln Met Glu Tyr Leu Asp Met
                325                 330                 335

Val Val Asn Glu Thr Leu Arg Leu Phe Pro Ile Ala Met Arg Leu Glu
            340                 345                 350

Arg Val Cys Lys Lys Asp Val Glu Ile Asn Gly Met Phe Ile Pro Lys
        355                 360                 365

Gly Val Val Val Met Ile Pro Ser Tyr Ala Leu His Arg Asp Pro Lys
    370                 375                 380

Tyr Trp Thr Glu Pro Glu Lys Phe Leu Pro Glu Arg Phe Ser Lys Lys
385                 390                 395                 400

Asn Lys Asp Asn Ile Asp Pro Tyr Ile Tyr Thr Pro Phe Gly Ser Gly
                405                 410                 415

Pro Arg Asn Cys Ile Gly Met Arg Phe Ala Leu Met Asn Met Lys Leu
            420                 425                 430

Ala Leu Ile Arg Val Leu Gln Asn Phe Ser Phe Lys Pro Cys Lys Glu
        435                 440                 445

Thr Gln Ile Pro Leu Lys Leu Ser Leu Gly Gly Leu Leu Gln Pro Glu
    450                 455                 460

Lys Pro Val Val Leu Lys Val Glu Ser Arg Asp Gly Thr Val Ser Gly
465                 470                 475                 480
```

Ala His His His His
485

<210> 9
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Motif

<400> 9
Ala Lys Lys Thr Ser Ser Lys Gly Arg
1 5

<210> 10
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Leader sequence

<400> 10
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg
1 5 10

<210> 11
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: N-terminal sequence of 3A4

<400> 11
Met Ala Tyr Gly Thr His Ser His Gly Leu Phe Lys Lys
1 5 10

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligonucleotide

<400> 12
caagaggaag cccgctgcct tgtggaggag 30

<210> 13
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 13
ctcctccaca aggcagcggg cttcctcttg 30

<210> 14
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 14
ccctcgtcac tttctggatg aaggtggaaa ttttaag 37

<210> 15
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide

<400> 15
cttaaaattt ccaccttcat ccagaaagtg acgaggg 37

<210> 16
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 16
catatggcta aaaaaacctc ttctaaaggc cgaccgccgg gtccgctgcc 50

<210> 17
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 17
gctgccaggc ctgggtaacc tgctgcatgt ggacttccag aacaccccgt 50

<210> 18
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 18
actgcttcga ccagctgcgt cgtcgtttcg gtgacgtgtt ctctctgcag 50

<210> 19
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 19
ctggcttgga ccccggttgt tgttctgaac ggtctggctg ctgttcgcga 50

<210> 20
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 20
agctctggtt acccacggtg aagacaccgc tgaccgtccg ccggtcccga 50

<210> 21
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 21
tcacccagat cctgggtttt ggtccgcgtt cccaaggtgt tttcctggct 50

<210> 22
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 22
cgttacggac cggcttggcg tgaacagcgt cgtttctctg tttctaccct 50

<210> 23
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 23
gcgtaacctg ggtctgggta aaaaatctct ggaacagtgg gttaccgaag 50

<210> 24
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 24
aagctgcatg cctgtgcgct gctttcgcta accactctgg tcgtccgttc 50

<210> 25
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 25
cgtccgaacg gtctgctgga caaagctgtt tctaacgtta tcgcttctct 50

<210> 26

<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 26
gacctgcggc cgccgtttcg aatacgacga cccgcgtttc ctgcgtctgc 50

<210> 27
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 27
tggacctggc tcaggaaggt ctgaaagagg agtctggttt cctgcgtgaa 50

<210> 28
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 28
gttctgaacg ctgttccggt tctgctgcac atcccagctc tggctggtaa 50

<210> 29
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 29
agttctgcgt ttccagaaag cattcctgac ccagctggac gaactgctga 50

<210> 30
<211> 50
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 30
ccgaacaccg tatgacctgg gacccggctc agccgccacg tgacctgacc 50

<210> 31
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 31
gaagctttcc tggctgaaat ggaaaaagct aaaggtaacc cggaatcttc 50

<210> 32
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 32
tttcaacgat gaaaatctgc gtatcgttgt tgctgacctg ttctccgcgg 50

<210> 33
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 33
gtatggttac cacctctacc accctggctt ggggtctgct gctgatgatc 50

<210> 34
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 34
ctgcacccgg atgtacagcg tcgtgttcag caggaaatcg acgacgttat 50

<210> 35
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 35
tggccaggtt cgtcggccgg aaatgggtga ccaggctcac atgccgtaca 50

<210> 36
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 36
ccaccgctgt tatccacgaa gttcagcgct tcggtgacat cgttccgctg 50

<210> 37
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 37
ggtatgaccc acatgacctc tcgtgacatc gaagttcagg gtttccgtat 50

<210> 38
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 38
cccgaaaggt accaccctga tcaccaacct gtcttctgtt ctgaaagacg 50

<210> 39
<211> 50
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 39
aagctgtttg ggaaaaaccg ttccgtttcc atccggaaca cttcctggac 50

<210> 40
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 40
gctcagggtc acttcgttaa accggaagct ttcctgccgt tctctgctgg 50

<210> 41
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 41
tcgtcgtgct tgcctgggtg aaccgctggc tcgtatggaa ctgttcctgt 50

<210> 42
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 42
tcttcacctc tctgctgcag cacttctctt tctctgttcc gaccggtcag 50

<210> 43
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 43
ccgcgtccgt ctcaccacgg tgttttcgct ttcctggttt ctccgtctcc 50

<210> 44
<211> 77
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 44
gtcgactcag tggtggtggt gagctccacg cggaacagcg cacagttcgt acggagacgg 60
agaaaccagg aaagcga 77

<210> 45
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 45
aaacaccgtg gtgagacgga cgcggctgac cggtcggaac agagaaagag 50

<210> 46
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 46
aagtgctgca gcagagaggt gaagaacagg aacagttcca tacgagccag 50

<210> 47
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 47
cggttcaccc aggcaagcac gacgaccagc agagaacggc aggaaagctt 50

<210> 48
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 48
ccggtttaac gaagtgaccc tgagcgtcca ggaagtgttc cggatggaaa 50

<210> 49
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 49
cggaacggtt tttcccaaac agcttcgtct ttcagaacag aagacaggtt 50

<210> 50
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 50
ggtgatcagg gtggtacctt tcgggatacg gaaaccctga acttcgatgt 50

<210> 51
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 51
cacgagaggt catgtgggtc atacccagcg gaacgatgtc accgaagcgc 50

<210> 52
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 52
tgaacttcgt ggataacagc ggtggtgtac ggcatgtgag cctggtcacc 50

<210> 53
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 53
catttccggc cgacgaacct ggccaataac gtcgtcgatt tcctgctgaa 50

<210> 54
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 54
cacgacgctg tacatccggg tgcaggatca tcagcagcag accccaagcc 50

<210> 55
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 55
agggtggtag aggtggtaac catacccgcg gagaacaggt cagcaacaac 50

<210> 56
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 56
gatacgcaga ttttcatcgt tgaaagaaga ttccgggtta cctttagctt 50

<210> 57
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for the 2D6 assembly

<400> 57
tttccatttc agccaggaaa gcttcggtca ggtcacgtgg cggctgagcc 50

<210> 58
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for the 2D6 assembly

<400> 58
gggtcccagg tcatacggtg ttcggtcagc agttcgtcca gctgggtcag 50

<210> 59
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for the 2D6 assembly

<400> 59
gaatgctttc tggaaacgca gaactttacc agccagagct gggatgtgca 50

<210> 60
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for the 2D6 assembly

<400> 60
gcagaaccgg aacagcgttc agaacttcac gcaggaaacc agactcctct 50

<210> 61

<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 61
ttcagacctt cctgagccag gtccagcaga cgcaggaaac gcgggtcgtc 50

<210> 62
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 62
gtattcgaaa cggcggccgc aggtcagaga agcgataacg ttagaaacag 50

<210> 63
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 63
ctttgtccag cagaccgttc ggacggaacg gacgaccaga gtggttagcg 50

<210> 64
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 64
aaagcagcgc acaggcatgc agcttcttcg gtaacccact gttccagaga 50

<210> 65
<211> 50
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 65
ttttttaccc agacccaggt tacgcagggt agaaacagag aaacgacgct 50

<210> 66
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 66
gttcacgcca agccggtccg taacgagcca ggaaaacacc ttgggaacgc 50

<210> 67
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 67
ggaccaaaac ccaggatctg ggtgatcggg accggcggac ggtcagcggt 50

<210> 68
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 68
gtcttcaccg tgggtaacca gagcttcgcg aacagcagcc agaccgttca 50

<210> 69
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 69
gaacaacaac cggggtccaa gccagctgca gagagaacac gtcaccgaaa 50

<210> 70
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 70
cgacgacgca gctggtcgaa gcagtacggg gtgttctgga agtccacatg 50

<210> 71
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Oligo used for
the 2D6 assembly

<400> 71
cagcaggtta cccaggcctg gcagcggcag cggacccggc ggtcggcctt 50

<210> 72
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 72
gtaacctggg tctgggtaaa aaatctctg 29

<210> 73
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 73
cagagatttt ttacccagac ccaggttac 29

<210> 74
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 74
ggaattcata tggctctcat cccagacttg gcc 33

<210> 75
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 75
tgcggtcgac tcaatggtga tggtgggctc cacttacggt gccatcc 47

<210> 76
<211> 69
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 76
ttaacatatg gcatatggta ctcattcaca tggtctgttt aaaaaactgg gaattccagg 60
gcccacacc 69

<210> 77
<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9-FGloop

<400> 77
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
1 5 10 15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys Asp Ile Ser
20 25 30

Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val Phe Thr Leu
35 40 45

Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr Glu Ala Val
50 55 60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile
65 70 75 80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser

```
                85                  90                  95

Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
            100                 105                 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
        115                 120                 125

Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
    130                 135                 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145                 150                 155                 160

Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
                165                 170                 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu Ser Ser Pro
            180                 185                 190

Trp Ile Gln Val Tyr Asn Asn Phe Pro Ala Leu Leu Asp Tyr Phe Pro
        195                 200                 205

Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met Lys Ser Tyr
    210                 215                 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Met Asn Asn
225                 230                 235                 240

Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu Lys Glu Lys
                245                 250                 255

His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu Asn Thr Ala
            260                 265                 270

Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
        275                 280                 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
    290                 295                 300

Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser Pro Cys Met
305                 310                 315                 320

Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
                325                 330                 335

Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His Ala Val Thr
            340                 345                 350

Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
        355                 360                 365

Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
    370                 375                 380

Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly Gly Asn Phe
```

385 390 395 400

Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe Leu Thr Ser
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro Lys Asn Leu
435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

Gln Leu Cys Phe Ile Pro Val His His His His
465 470 475

<210> 78
<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9-P220 V60C

<400> 78
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
1 5 10 15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys Asp Ile Ser
20 25 30

Lys Ser Leu Thr Asn Leu Ser Lys Cys Tyr Gly Pro Val Phe Thr Leu
35 40 45

Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr Glu Ala Val
50 55 60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile
65 70 75 80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser
85 90 95

Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
100 105 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
115 120 125

Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
130 135 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145 150 155 160

Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
165 170 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu Ser Ser Pro
180 185 190

Trp Ile Gln Ile Cys Asn Asn Phe Pro Thr Ile Ile Asp Tyr Phe Pro
195 200 205

Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met Lys Ser Tyr
210 215 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Met Asn Asn
225 230 235 240

Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu Lys Glu Lys
245 250 255

His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu Asn Thr Ala
260 265 270

Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
275 280 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
290 295 300

Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser Pro Cys Met
305 310 315 320

Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
325 330 335

Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His Ala Val Thr
340 345 350

Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
355 360 365

Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
370 375 380

Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly Gly Asn Phe
385 390 395 400

Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe Leu Thr Ser
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro Lys Asn Leu
435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

```
Gln Leu Cys Phe Ile Pro Val His His His His
465                 470                 475


<210> 79
<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9-P220

<400> 79
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
  1               5                  10                  15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys Asp Ile Ser
             20                  25                  30

Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val Phe Thr Leu
         35                  40                  45

Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr Glu Ala Val
     50                  55                  60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile
 65                  70                  75                  80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser
                 85                  90                  95

Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
            100                 105                 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
        115                 120                 125

Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
    130                 135                 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145                 150                 155                 160

Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
                165                 170                 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu Ser Ser Pro
            180                 185                 190

Trp Ile Gln Ile Cys Asn Asn Phe Pro Thr Ile Ile Asp Tyr Phe Pro
        195                 200                 205

Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met Lys Ser Tyr
    210                 215                 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Met Asn Asn
225                 230                 235                 240
```

Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu Lys Glu Lys
245 250 255

His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu Asn Thr Ala
260 265 270

Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
275 280 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
290 295 300

Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser Pro Cys Met
305 310 315 320

Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
325 330 335

Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His Ala Val Thr
340 345 350

Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
355 360 365

Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
370 375 380

Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly Gly Asn Phe
385 390 395 400

Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe Leu Thr Ser
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro Lys Asn Leu
435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

Gln Leu Cys Phe Ile Pro Val His His His His
465 470 475

<210> 80
<211> 475
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9-FG
Loop-K206E

<400> 80

```
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
 1               5                  10                  15

Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys Asp Ile Ser
            20                  25                  30

Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val Phe Thr Leu
        35                  40                  45

Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr Glu Ala Val
    50                  55                  60

Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile
65                  70                  75                  80

Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile Val Phe Ser
                85                  90                  95

Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr Leu
            100                 105                 110

Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg Val Gln Glu
        115                 120                 125

Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys Ala Ser Pro
    130                 135                 140

Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn Val Ile Cys
145                 150                 155                 160

Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln Gln Phe Leu
                165                 170                 175

Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Glu Ile Leu Ser Ser Pro
            180                 185                 190

Trp Ile Gln Val Tyr Asn Asn Phe Pro Ala Leu Leu Asp Tyr Phe Pro
        195                 200                 205

Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met Lys Ser Tyr
    210                 215                 220

Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp Met Asn Asn
225                 230                 235                 240

Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu Lys Glu Lys
                245                 250                 255

His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu Asn Thr Ala
            260                 265                 270

Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr Thr Leu Arg
        275                 280                 285

Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr Ala Lys Val
    290                 295                 300
```

Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser Pro Cys Met
305 310 315 320

Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val His Glu Val
325 330 335

Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His Ala Val Thr
340 345 350

Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly Thr Thr Ile
355 360 365

Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu Phe Pro Asn
370 375 380

Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly Gly Asn Phe
385 390 395 400

Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys Arg Ile Cys
405 410 415

Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe Leu Thr Ser
420 425 430

Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro Lys Asn Leu
435 440 445

Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro Pro Phe Tyr
450 455 460

Gln Leu Cys Phe Ile Pro Val His His His His
465 470 475

<210> 81
<211> 494
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9 wild type
P450

<400> 81
Met Asp Ser Leu Val Val Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
1 5 10 15

Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
20 25 30

Pro Thr Pro Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys
35 40 45

Asp Ile Ser Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val
50 55 60

Phe Thr Leu Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr

```
65                  70                  75                  80

Glu Ala Val Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly
                85                  90                  95

Arg Gly Ile Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile
            100                 105                 110

Val Phe Ser Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu
        115                 120                 125

Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg
    130                 135                 140

Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
145                 150                 155                 160

Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
                165                 170                 175

Val Ile Cys Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln
            180                 185                 190

Gln Phe Leu Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu
        195                 200                 205

Ser Ser Pro Trp Ile Gln Ile Cys Asn Asn Phe Ser Pro Ile Ile Asp
    210                 215                 220

Tyr Phe Pro Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met
225                 230                 235                 240

Lys Ser Tyr Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp
                245                 250                 255

Met Asn Asn Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu
            260                 265                 270

Lys Glu Lys His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu
        275                 280                 285

Asn Thr Ala Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr
    290                 295                 300

Thr Leu Arg Tyr Ala Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr
305                 310                 315                 320

Ala Lys Val Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser
                325                 330                 335

Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val
            340                 345                 350

His Glu Val Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His
        355                 360                 365

Ala Val Thr Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly
```

```
    370                 375                 380

Thr Thr Ile Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu
385                 390                 395                 400

Phe Pro Asn Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly
                405                 410                 415

Gly Asn Phe Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys
            420                 425                 430

Arg Ile Cys Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe
        435                 440                 445

Leu Thr Ser Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro
    450                 455                 460

Lys Asn Leu Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro
465                 470                 475                 480

Pro Phe Tyr Gln Leu Cys Phe Ile Pro Val His His His His
                485                 490


<210> 82
<211> 40
<212> PRT
<213> Homo sapiens

<400> 82
Met Asp Ser Leu Val Val Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
  1               5                  10                  15

Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
             20                  25                  30

Pro Thr Pro Leu Pro Val Ile Gly
         35                  40


<210> 83
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 2C9 and 2C19
      truncation

<400> 83
Met Ala Lys Lys Thr Ser Ser Lys Gly Arg Pro Pro Gly Pro Thr Pro
  1               5                  10                  15

Leu Pro Val Ile Gly
             20
```

<210> 84
<211> 40
<212> PRT
<213> Homo sapiens

<400> 84
Met Asp Pro Phe Val Val Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
1 5 10 15

Leu Ser Ile Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
20 25 30

Pro Thr Pro Leu Pro Val Ile Gly
35 40